ACTIVATION DETECTION IN FUNCTIONAL MRI BASED ON NON-SEPARABLE SPACE-TIME NOISE MODELS

Joonki Noh[†] *and Victor Solo*[‡]

[†]Dept.of Electrical Eng. and Computer Science, University of Michigan, Ann Arbor, USA,
[‡]School of Electrical Eng., University of New South Wales, Sydney, AUSTRALIA.

ABSTRACT

Detecting activated regions in the human brain by cognitive tasks is a significant task in the data analysis using functional MRI (FMRI). To create a detection statistic for activation, noise models under two assumptions; 1) spatial independence and 2) space-time separability have been dominantly used in the FMRI data analysis. In this paper, we propose a novel detection statistic derived from noise models with spatiotemporal correlation and without space-time separability. In order to obtain a sufficiently flexible class of noise models for non-separable space-time processes, an unusual noise modeling based on truncated cepstrum expansion is suggested. Developed methods are applied to a human dataset.

Index Terms— Functional MRI, activation detection, non-separable space-time noise models, spatiotemporal correlation, and the parametric cepstrum

1. INTRODUCTION

The advent of brain imaging techniques allows researchers in many fields to study the healthy and living human subject while the subject is performing cognitive tasks without the need for surgery. Among non-invasive techniques, FMRI is probably one of the most popular tool to analyze and visualize the brain activity and a vast amount of statistical methods for FMRI data analysis have been proposed. These can be categorized into two groups; univoxel methods mainly based on classical time series analysis and multivoxel methods based on multivariate analysis, e.g., principal component analysis (PCA) and independent component analysis (ICA).

With rare exceptions, most previous methods have been based on two main assumptions on noise processes in the human brain; spatial independence and space-time separability. In the activation study, assuming space-time separability implies that pure spatial and pure temporal operations can be separately and sequentially done to properly detect activations in the brain induced by given stimuli. For example in SPM, spatial smoothing by Gaussian kernel (pure spatial operation) and GLM (involving pure temporal operation) are performed sequentially to construct a detection statistic such as t-statistic or F-statistic. [1] is an exception in previous approaches. However, [1] suggested heavily constrained noise models combining temporal AR(p) and spatial AR(1) processes and did not provided a detection statistic for activation from those noise models.

We recently developed a detection statistic from noise models fully considering spatial and temporal correlations under a space-time separability assumption [2]. By extending [2], we develop a detection statistic from non-separable space-time noise models in this paper. The detection statistic provided in [2] is a special case of the one developed in this paper.

2. FMRI MEASUREMENT MODEL

We consider a real-valued measurement model that has the following additive form at a time t and a voxel v:

$$y_{t,v} = d_{t,v} + s_{t,v} + w_{t,v}, \tag{1}$$

where $d_{t,v}$ models nuisance signal component, e.g., slowly varying temporal drift. It is found that linear drift is reasonably adequate [3]. $s_{t,v}$ that is signal of our main interest denotes the blood oxygenation level dependent (BOLD) response induced by given stimuli. $w_{t,v}$ is non-separable random noise that is spatiotemporally correlated and stationary. This noise has several sources; 1) MR scanner and 2) background processes from physiologically unknown origins in the brain. The noise from 1) is modeled as a white process and the noise from 2) can be modeled as a spatiotemporal colored process. For integer-valued t and v, we assume that $y_{t,v}$ is observed from a rectangle, $\{0,\ldots,T-1\} \times \{0,\ldots,M-1\}$. More details about the signal and noise models are now discussed.

2.1. Signal Model Formulation

The BOLD response can be linearly modeled as follows:

$$s_{t,v} = h_t^T f_v * c_t \triangleq \xi_t^T f_v, \tag{2}$$

where $h_t \triangleq [h_{1,t},\ldots,h_{L,t}]^T$ is a set of basis functions and $f_v \triangleq [f_{1,v},\ldots,f_{L,v}]^T$ is the associated activation amplitude. The given stimulus is denoted as c_t and L is the number of basis functions for the BOLD response modeling. (2) covers various modelings of the BOLD response; the parametric modeling with the canonical hemodynamic response function, FIR modeling, and Laguerre modeling [4]. The choice of basis functions is crucial for a good modeling of the BOLD response with a small amount of computations. The FIR modeling typically requires a high order FIR filter, whereas the Laguerre polynomials provide a compact and exact modeling of the BOLD response [4].

2.2. Noise Model Formulation

For modeling the non-separable and spatiotemporally colored noise $w_{t,v}$ in (1), we need a class of noise models that is flexible enough to handle such processes interesting in FMRI. Our approach to noise modeling is based on a cepstrum expansion, called the parametric cepstrum, which allows several advantages over classical AR-based methods [5]. For example, the parametric cepstrum covers a wide range of non-separable spatiotemporal processes and requires nearly linear model fitting by FFTs [2, 5]. However, 2D or 3D AR-based methods require highly nonlinear model fitting, as a matter of fact, computationally prohibitive.

The parametric cepstrum can be defined by truncating a Fourier series expansion of the logarithm of power spectral density (PSD). For $\omega_k \triangleq 2\pi k/T$ and $\lambda_l \triangleq 2\pi l/M$,

$$\log F_{k,l} \;=\; \sum_{t=-n}^{n} \sum_{v=-p}^{p} \theta_{t,v} e^{-j(\omega_k t + \lambda_l v)}, \tag{3}$$

where $F_{k,l}$ is PSD and (n, p) denotes noise model orders. $\theta_{t,v}$'s are called cepstral coefficients. Thus the parametric cepstrum allows a parametric modeling of PSD. Details of noise model fitting with the parametric cepstrum are not given in this paper due to space limit. The reader who is interested in more details is referred to [5] and [6, Chap.4].

2.3. Measurement Model Reformulation

From (2) and (3), we have the following fully parameterized formulation for FMRI measurement:

$$y_{t,v} = X_t^T \beta_v + \xi_t^T f_v + w_{t,v}, \tag{4}$$

where the PSD of $w_{t,v}$ is given by $F_{k,l}$ and the linear drift $d_{t,v}$ is modeled by $X_t^T \triangleq [1, t]$ and $\beta_v^T \triangleq [m_v, b_v]$.

We emphasize two important points that produce big differences between the model in (4) and a conventional model involving spatial smoothing by Gaussian kernel (SSK). First note that the activation amplitude f_v in (4) is a parameter we need to estimate "voxel by voxel" from collected FMRI data. However, the activation amplitude in the conventional model is assumed to be a Gaussian point spread function (PSF) imposing spatial continuity of activations [7]. Based on Gaussian PSF and under a spatial white noise assumption, [8] showed that likelihood ratio test leaded to a matched filter involving SSK. Secondly, the noise $w_{t,v}$ is a spatiotemporally colored process that is not under a space-time separability assumption, whereas the noise in the conventional model is under spatial independence and space-time separability. For temporal noise modeling, temporal AR processes have been used in the conventional approach.

3. DETECTION STATISTIC

3.1. General Development

The consideration of the spatiotemporal structure of $w_{t,v}$ in the temporal frequency and spatial wave-number domains simplifies the discussion. Taking spatiotemporal DFTs in (4) gives us, for $k(= 0, \ldots, T-1)$ and $l(= 0, \ldots, M-1)$,

$$\tilde{y}_{k,l} = \tilde{X}_k^T \tilde{\beta}_l + \tilde{\xi}_k^T \tilde{f}_l + \tilde{w}_{k,l}, \tag{5}$$

where, e.g., $\tilde{\xi}_k$ denotes the temporal DFT of ξ_t and $\tilde{f}_l$ is the spatial DFT of f_v. For large T and M, under spatiotemporal stationarity and some regularity conditions involving joint cumulants, $\tilde{w}_{k,l}$ obeys central limit theorem (CLT) [9] leading to, for $\forall (k,l) \neq (0,0)$,

$$\frac{1}{\sqrt{TM}}\, \tilde{w}_{k,l} \;\sim\; \mathcal{N}_c(0, F_{k,l}) \tag{6}$$

and $\tilde{w}_{0,0} \sim \mathcal{N}(0, TM \cdot F_{0,0})$, where $\mathcal{N}_c$ and $\mathcal{N}$ are complex and real Gaussian distributions, respectively. In addition, $\{\tilde{w}_{k,l}\}_{l=0,\ldots,B}^{k=0,\ldots,A}$ are jointly Gaussian and independent for any (A, B) such that $\{1 \leq A \leq (T-1)/2, 0 \leq B \leq M-1;\ A = 0, 1 \leq B \leq (M-1)/2\}$, where T and M are assumed odd.

To develop a detection statistic for a voxel location of interest, we consider a hypothesis testing problem for the whole region of interest (ROI), namely

$$\begin{array}{lcll} H_0 & : & f_v = 0 & for\ all \quad v, \\ H_1 & : & f_v \neq 0 & for\ some \quad v. \end{array} \tag{7}$$

In the temporal frequency and spatial wave-number domains, we have an equivalent hypothesis testing problem;

$$\begin{array}{lcll} H_0 & : & \tilde{f}_l = 0 & for\ all \quad l, \\ H_1 & : & \tilde{f}_l \neq 0 & for\ some \quad l, \end{array} \tag{8}$$

where H_0 states there is no activation in the ROI and H_1 states its alternative. Based on (6) and (8), we construct a likelihood ratio test (LRT) statistic $\mathbf{\Lambda}$ whose logarithm defining $LRT(\triangleq 2\log \mathbf{\Lambda})$ is given by

$$\begin{aligned} -2\log \mathbf{\Lambda} \;&=\; \sum_{k=0}^{T-1} \sum_{l=0}^{M-1} \log \widehat{F}_{1,k,l} - \log \widehat{F}_{0,k,l} \\ &+\; \sum_{k=0}^{T-1} \sum_{l=0}^{M-1} \frac{|\tilde{e}_{1,k,l}|^2}{TM \cdot \widehat{F}_{1,k,l}} - \frac{|\tilde{e}_{0,k,l}|^2}{TM \cdot \widehat{F}_{0,k,l}}, \end{aligned} \tag{9}$$

where all estimates are MLEs, e.g., $\widehat{F}_{1,k,l}$ is the MLE of $F_{k,l}$ under H_1. Residuals $\tilde{e}_{k,l}$ are defined as $\tilde{e}_{0,k,l} \triangleq \tilde{y}_{k,l} - \tilde{X}_k^T \widehat{\tilde{\beta}}_{0,l}$ under H_0 and $\tilde{e}_{1,k,l} \triangleq \tilde{y}_{k,l} - \tilde{X}_k^T \widehat{\tilde{\beta}}_{1,l} - \tilde{\xi}_k^T \widehat{\tilde{f}}_l$ under H_1.

For constructing LRT in (9), we use an approximate likelihood derived from (6) in the frequency domain. This construction of an approximate likelihood for a parametric model in the frequency domain can work well even for small sample sizes and thus widely used in times series and econometrics. In practice, the validity of the approximate likelihood from the frequency domain approach can be investigated by statistical diagnoses, e.g., QQ-plot.

3.2. Spatial Decomposition

To develop a voxel-wise detection statistic, we manipulate LRT from (9) in two ways. We firstly have an equivalent expression to LRT in the space-time domain via Parseval's relation. We secondly define LRT_v by performing spatial decomposition of the LRT in the space-time domain, i.e., $LRT = \sum_v LRT_v$. It can be shown that this spatial decomposition of LRT produces the following spatio-temporal LRT (ST-LRT) in two pieces; a noise piece and signal piece,

$$LRT_v \;\triangleq\; LRT_v^N + LRT_v^S, \tag{10}$$

where LRT_v^N reflects the difference between the estimates of PSD under H_0 and H_1. The noise and signal pieces are defined as

$$\begin{aligned} LRT_v^N \;&\triangleq\; T\left(\widehat{\theta}_{0,0,0} - \widehat{\theta}_{1,0,0}\right) + \sum_{t=0}^{T-1} \left(\varepsilon_{0,t,v}^2 - \varepsilon_{1,t,v}^2\right), \\ LRT_v^S \;&\triangleq\; \mathbf{s}_{y,v}^H \mathbf{s}_{y,v}, \end{aligned} \tag{11}$$

and

$$\mathbf{s}_{y,v} \overset{DFT}{\longleftrightarrow} \tilde{\mathbf{s}}_{y,l}, \qquad \tilde{\mathbf{s}}_{y,l} \triangleq \mathbf{S}_{\xi\xi,l}^{-\frac{1}{2}} \tilde{\mathbf{s}}_{y\xi,l},$$

where $\widehat{\theta}_{j,0,0}$ is the MLE of $\theta_{0,0}$ (cepstral coefficient at the origin) and $\varepsilon_{j,t,v}$ is spatiotemporally whitened data $y_{j,t,v}$ under H_j for $j = 0, 1$, i.e.,

$$\varepsilon_{j,t,v} \triangleq (g_{j,t,v} \circledast \circledast y_{j,t,v}). \tag{12}$$

$g_{j,t,v}$ is a spatiotemporal whitening filter that is temporally causal and $y_{j,t,v}$ is adjusted from $y_{t,v}$ by removing out nuisance signal component estimate. For the signal piece LRT_v^S, $\tilde{\mathbf{s}}_{y\xi,l}$ and $\mathbf{S}_{\xi\xi,l}$ are defined as

$$\tilde{\mathbf{s}}_{y\xi,l} \triangleq \frac{1}{T}\cdot\sum_{k=0}^{T-1}\frac{\tilde{y}^*_{1,k,l}\tilde{\xi}_{1,k,l}}{\widehat{F}_{1,k,l}},\quad \mathbf{S}_{\xi\xi,l} \triangleq \frac{1}{T}\cdot\sum_{k=0}^{T-1}\frac{\tilde{\xi}_{1,k,l}\cdot\tilde{\xi}^H_{1,k,l}}{\widehat{F}_{1,k,l}}, \tag{13}$$

where $\tilde{\xi}_{1,k,l}$ is adjusted from $\tilde{\xi}_k$ by removing out nuisance signal component estimate and is dependent on (k,l) due to the non-separable nature of noise processes. We have the following relation between the spatiotemporal whitening filter $g_{t,v}$ and the PSD $F_{k,l}$: Under H_j for $j=0,1$,

$$g_{j,t,v} \overset{DFT}{\longleftrightarrow} \tilde{g}_{j,k,l}, \qquad |\tilde{g}_{j,k,l}|^2 = \frac{1}{\widehat{F}_{j,k,l}}, \tag{14}$$

where $g_{j,t,v}$ is temporally causal. In (13), $\tilde{\mathbf{s}}_{y\xi,l}$ evaluates the cross correlation between $y_{t,v}$ (data) and $\xi_{t,v}$ (temporal response) after the spatiotemporal whitening with $g_{t,v}$ is done in the frequency domain. Similarly, $\mathbf{S}_{\xi\xi,l}$ computes the auto-correlation of spatiotemporally whitened temporal response. Details of the derivation of (10) from (9) are provided in [6, Chap.3]. If a space-time separability of noise process is assumed, the signal piece of ST-LRT in (11) reduces to a closed form in the space-time domain and $g_{t,v}$ is decomposed into its pure temporal and pure spatial pieces [2].

We here emphasize three significant differences between the conventional GLM involving SSK (called SSK-GLM in this paper.) and the developed ST-LRT. Firstly, the spatiotemporal whitening filter $g_{t,v}$ is based on spatiotemporal correlation without space-time separability, whereas SSK-GLM is established on the basis of spatial independence and space-time separability as we outlined in several places of this paper. Secondly, different whitening filters ($g_{0,t,v}$ and $g_{1,t,v}$) are estimated from and applied to data under H_0 and H_1 for ST-LRT. However, spatial kernel smoothing with the same FWHM is performed in SSK-GLM regardless of hypothesis. Finally, it can be shown that the spatiotemporal whitening filter $g_{t,v}$ is more like a differentiator that is opposite to a smoother used in SSK-GLM.

3.3. Family-Wise Error Rate Control

By computing ST-LRTs over all voxels in a ROI and thresholding them with a pre-determined cutoff point, a thresholded detection statistic called an activation map can be obtained. A threshold is determined for the purpose of controlling an overall error rate. Since we usually have many voxels in a ROI, that is a multiple comparison problem. One widely used measure is family-wise error (FWE) rate whose definition is given under H_0 as follows:

$$\text{FWE} \triangleq Pr\left(\bigcup_{v=1}^{M}\{LRT_v > \gamma\}\right) = Pr\left(\max_v LRT_v > \gamma\right), \tag{15}$$

where threshold γ is determined for a significance level α, typically 0.05 in FMRI. False discovery rate is an alternative choice to FWE.

From (15), we need an approximate distribution of the maximum of LRT_v to determine γ for a given α. If β_v (nuisance signal) and $F_{k,l}$ (PSD) are known, we have the following asymptotic properties under H_0 that are useful to determine γ:

(P1) $\varepsilon_{0,t,v} \triangleq (g_{0,t,v} \circledast\circledast y_{0,t,v})$ is a spatiotemporally white noise that obeys $\mathcal{N}(0,1)$ for any (t,v),

(P2) $LRT_u \overset{i.d.}{\sim} LRT_v$ for all $u \neq v$,

(P3) $LRT_v \sim \chi^2_L$,

where $x \overset{i.d.}{\sim} y$ means x and y are independently distributed and χ^2_L denotes a chi-square distribution with L degrees of freedom. Proofs of (P1)-(P3) are given in [6, Chap.3]. Due to (P2)-(P3), threshold γ can be analytically determined. For a significance level α, we have

$$\gamma(\alpha) = \Psi(\chi^2_L)^{-1}\left(\sqrt[M]{1-\alpha}\right), \tag{16}$$

where $\Psi(\chi^2_L)$ denotes the cumulative density function of χ^2_L. Thus, no random field theory (RFT) is required to control FWE rate for ST-LRT.

4. MODEL SELECTION AND COMPARISON

By means of Akaike information criterion (AIC), we can not only compare two models with different structures, e.g., ST-LRT model and SSK-GLM, but also models with the same structure to determine proper orders, e.g, (L,n,p) in (4) for ST-LRT model. Recall that L denotes the number of basis functions to model the BOLD response, thus L is the number of interesting signal parameters. (n,p) contains the orders for modeling PSD in (3), thus determines the number of noise parameters. We first consider a model selection for ST-LRT model to determine (L,n,p), then move the discussion to a model comparison of ST-LRT model and SSK-GLM. A similar approach in FMRI can be found in [3].

From the discussion in Section 3, we have the following AIC:

$$AIC_{L,R} \triangleq \sum_{k=0}^{T-1}\sum_{l=0}^{M-1}\log\widehat{F}_{1,k,l} + \frac{|\tilde{e}_{1,k,l}|^2}{TM\cdot\widehat{F}_{1,k,l}} + 2(L+R), \tag{17}$$

where the first two terms denote the negative approximate likelihood under H_1 as in (9) and the last term including the number of cepstral coefficients R ($\triangleq 2np+n+p+1$) means the model complexity. Minimizing (17) produces the model orders (L,n,p) for the fully parameterized model in (4). Alternatives to AIC can be Bayesian information criterion (BIC) and minimum description length (MDL) that pick smaller models than does AIC.

For model comparison, the same ideas used to construct ST-LRT is applied to AIC. Applying Parseval's relation and then performing spatial decomposition of (17) yield an AIC map for ST-LRT model given by

$$AIC_v^{LRT} \triangleq T\cdot\widehat{\theta}_{1,0,0} + \sum_{t=0}^{T-1}(g_{1,t,v} \circledast\circledast e_{1,t,v})^2 + \frac{2(L+R)}{M}, \tag{18}$$

where $\circledast\circledast$ denotes spatiotemporal circular convolution and $e_{1,t,v}$ is the inverse DFT of residual $\tilde{e}_{1,k,l}$. Here, AIC_v^{LRT} represents the voxel-wise contribution to AIC for the whole ROI.

An AIC map for SSK-GLM can be also built up. It can be shown [6, Chap.4] that an AIC map for SSK-GLM is given by

$$AIC_v^{DLM} \triangleq \sum_{k=0}^{T-1}\log\widehat{\breve{F}}_k - 2T\,\theta_0^G + \sum_{t=0}^{T-1}\bar{e}^2_{t,v} + \frac{2(L+n_p)}{M}, \tag{19}$$

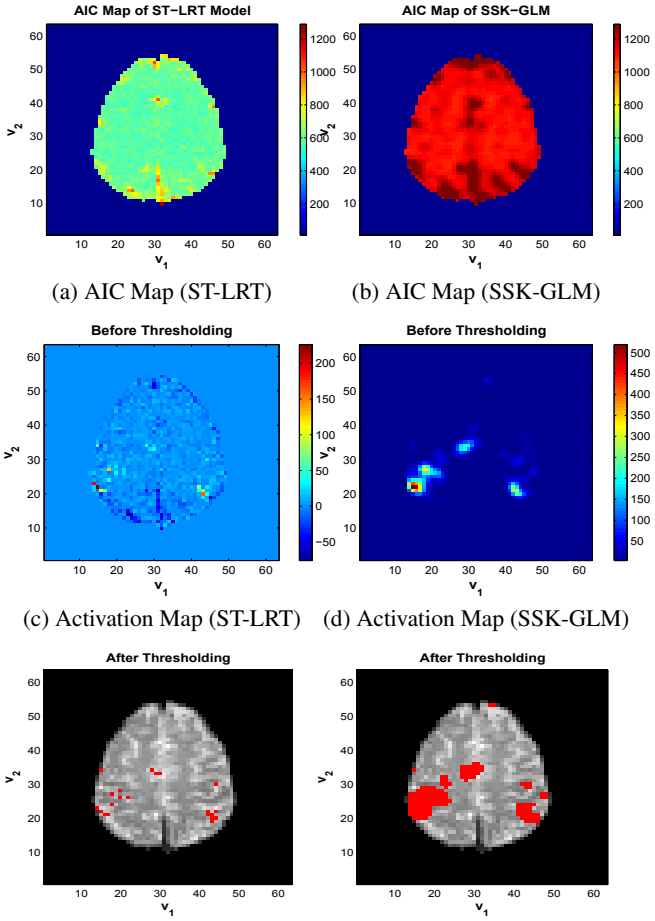
where $\breve{F}_k$ is pure temporal PSD of assumed noise models and θ_v^G is the cepstrum of a Gaussian amplitude kernel K_v^G for SSK. The residual is defined as $\bar{e}_{t,v} \triangleq \bar{y}_{t,v} - \bar{X}_t^T\widehat{\bar{\beta}}_v - \bar{\xi}_t^T\widehat{\bar{f}}_v$, where, e.g., $\bar{y}_{t,v}$ is the spatially smoothed and temporally filtered $y_{t,v}$ by K_v^G and ϕ_t, that is $\bar{y}_{t,v} \triangleq \phi_t * (K_v^G * y_{t,v})$. ϕ_t can be a temporal whitening filter or coloring filter. n_p is the number of parameters for AR-based noise modeling. For an AR(1) process as in SPM, we have $n_p = 2$ and $\breve{F}_k = \frac{\breve{\sigma}^2}{|1-\breve{\varphi}e^{-j\omega_k}|^2}$, where $\breve{\sigma}^2$ is the noise variance after SSK and $\breve{\varphi}$ is an AR(1) coefficient.

5. APPLICATION TO A HUMAN DATASET

The ST-LRT derived from non-separable space-time noise models is applied to a sample dataset from AFNI homepage (http://afni.nimh.nih.gov/afni/). This dataset was collected from an experiment where a human subject performed the right-hand finger-thumb opposition. The experiment was done in 3T MRI scanner and the repetition time is set to 2 seconds. For brevity, we pick a 2D axial slice where corresponding motor responses are expected. We have $T = 99$, $M = 63 \times 63$, and a voxel whose size is $3.125 \times 3.125 \times 5mm^3$. We use a spatial mask to remove signals outside the subject's brain and then apply a spatiotemporal taper (Tukey-Hanning window) to reduce bias caused by edge effects in 3D spectral estimation.

For ST-LRT, Laguerre functions whose orders are up to 3 are used for the BOLD response modeling [4]. To check the validity of the approximate likelihood in the frequency domain used to build up ST-LRT, a statistical diagnosis is done. From [9], if CLT works well for this dataset, $I_{k,l}/\widehat{F}_{k,l}$ obeys $\chi^2_2/2$, where $I_{k,l}$ is Periodogram. We have no significant deviation in the QQ-plot of $\log(I_{k,l}/\widehat{F}_{k,l})$ and a Gumbel distribution (a logarithm of $\chi^2_2/2$), showing that the approximate likelihood behind ST-LRT works well for this AFNI dataset. For SSK-GLM, the FWHM of Gaussian smoothing kernel is set to 2.5 times of the voxel size as recommended in SPM. After SSK, temporal whitening based on AR(1) processes is performed for each voxel. Then voxel-wise F-statistics are constructed.

For comparing two competing models, two AIC maps, one is from non-separable ST-LRT model and the other is from SSK-GLM, are provided on the top of Fig.1. The AIC values of ST-LRT model are substantially lower than those of SSK-GLM, indicating that the ST-LRT model is on average closer to the unknown underlying truth generating the collected dataset than SSK-GLM. On the middle of Fig.1, two activation maps before thresholding are shown. On the bottom of Fig.1, two thresholded activation maps where red spots show activated regions in the brain are given for significance level 0.05. The activation map from ST-LRT shows sharper and more well-defined activated areas than that from SSK-GLM in the primary and secondary motor cortex.



(a) AIC Map (ST-LRT) (b) AIC Map (SSK-GLM)

(c) Activation Map (ST-LRT) (d) Activation Map (SSK-GLM)

(e) Activation Map (ST-LRT) (f) Activation Map (SSK-GLM)

Fig. 1. AIC Maps and Activation Maps from ST-LRT (left column) and SSK-GLM (right column): (c) and (d) are activation maps before thresholding. (e) and (f) are activation maps after thresholding.

6. CONCLUSIONS

We developed a detection statistic for activation from non-separable space-time noise models, in which spatiotemporal correlation is fully considered without space-time separability. For implementing the developed ST-LRT, the parametric cepstrum allowing nearly linear fitting procedure was used. Applying the ST-LRT to a human dataset showed its superiority over the conventional GLM involving spatial smoothing by Gaussian kernel via AIC maps.

7. REFERENCES

[1] M.W. Woolrich, M. Jenkinson, J.M. Brady, and S.M. Smith, "Fully Bayesian spatio-temporal modeling of fMRI data," *IEEE Trans. on Medical Imaging*, vol. 23, No.2, pp. 213–231, Feb. 2004.

[2] J. Noh and V. Solo, "A true spatio-temporal test statistic for activation detection in fMRI by parametric cepstrum," *In Proc. IEEE ICASSP, Honolulu USA*, 2007.

[3] V. Solo, P. Purdon, R. Weisskoff, and E. Brown, "A signal estimation approach to functional MRI," *IEEE Trans. on Medical Imaging*, vol. 20(1), pp. 26–35, 2001.

[4] V. Solo, C.J. Long, E.N. Brown, E. Aminoff, M. Bar, and S. Saha, "fMRI signal modeling using Laguerre polynomials," *In Proc. IEEE Intl. Conf. on Image Processing, Singapore*, Oct. 2004.

[5] V. Solo, "Modeling of two-dimensional random fields by parametric cepstrum," *IEEE Trans. on Information Theory*, vol. IT-32, No.6, pp. 743–750, Nov. 1986.

[6] J. Noh, "True spatio-temporal detection and estimation for functional magnetic resonance imaging," *PhD Thesis, The University of Michigan*, 2007 (available at http://www.eecs.umich.edu/ fessler/student/).

[7] K. Worsley, A. Evans, S. Marett, and P. Neelin, "A three dimensional statistical analysis for CBF activation studies in human brain," *Journal of Cerebral Blood Flow and Metabolism*, vol. 12, pp. 900–918, 1992.

[8] D.O. Siegmund and K.J. Worsley, "Testing for a signal with unknown location and scale in a stationary Gaussian random field," *Annals of statistics*, vol. 23, pp. 608–639, 1995.

[9] D.R. Brillinger, "Time Series: Data Analysis and Theory," *San Francisco, CA: Holdenday*, 1981.